MDPI

Editorial

The Impact of Next Generation Sequencing in Cancer Research

Katia Nones * and **Ann-Marie Patch**

QIMR Berghofer Medical Research Institute, Brisbane, 4006 Queensland, Australia; Ann-Marie.Patch@qimrberghofer.edu.au

* Correspondence: Katia.Nones@qimrberghofer.edu.au

Received: 24 September 2020; Accepted: 1 October 2020; Published: 12 October 2020

Next generation sequencing (NGS) describes the technical revolution that enabled massively parallel sequencing of fragmented nucleic acids, thus making possible our current genomic understanding of cancers. It took an international consortium, using factory-like centers and high-throughput Sanger (terminator dye sequencing), more than 10 years to release the completed reference human genome in 2003. The generation of this high quality human genome reference and associated gene model annotation was essential for cancer genomics studies. However, the scale of resources required to generate the human genome reference highlighted the need for innovative sequencing approaches that would redefine the term "high-throughput sequencing". NGS is the term used for the new form of sequencing that emerged in the mid-2000s and is the product of competing technologies that led to increased speed, capacity, accuracy of sequencing machines and the progressive fall in the cost of sequencing per base pair. It is this sequencing revolution which enabled researchers to conceive of sequencing two samples for each individual cancer patient: one representing the inherited genome and the other the cancer genome in order to identify cancer specific genomic alterations. The ability to undertake genome-wide sequencing studies for thousands of cancer patients at base-pair resolution has had a profound impact on our understanding of the mutational landscape of cancers and the mechanisms underlying the disease.

Cancer is one of the leading cause of death worldwide, and it is well accepted that cells in the human body may acquire genomic changes (mutations). These mutations may result in uncontrolled cellular growth, suppression of cell-death signaling, promote angiogenesis amongst other key hallmark characteristics that lead to the development of cancers [1]. Over the last decade, two large consortia led the way in generating and analyzing NGS data from different cancer types. The Cancer Genome Atlas (TCGA) [2], predominantly using exome sequencing [3], and The International Cancer Genome Consortium (ICGC) [4], using whole genome sequencing, explored somatic mutations across thousands of cases from >30 cancer types. These consortia combined both sequencing approaches with transcriptional, methylation and protein analyses. Their cataloguing studies enabled a pan-cancer analysis [5] that explored the molecular characteristics of cancers, identifying recurring mutational patterns irrespective of the tissue of origin of the cancer. These studies generated a massive amount of knowledge about many cancer types, including the identification of (1) driver genes implicated in tumor development, (2) molecular cancer subtypes and (3) a mutational basis for targeted treatment options [6]. These consortia also have been instrumental in developing and providing methods and tools for the recording, analyses and sharing of multi-omic data on an international scale. One such analysis is the identification of mutational signatures, using information from hundreds to thousands of passenger mutations that might not cause cancer but can be a source of information about deficient cellular DNA damage repair mechanisms or patient exposure to carcinogens [7,8].

A major outcome of the NGS exploration of cancers has been the realization that high levels of inter-tumor and intra-tumor heterogeneity are key features of malignant cancers. Much of the current

research efforts focus on how to identify patients whose cancers will respond to targeted treatments and how to negate the development of treatment resistance [9]. NGS of multiple samples, separated either spatially or temporally, from individual patients, have allowed the phylogenetic reconstruction of the evolution of mutational patterns and driver mutations arising in clonal compartments of cancers [10,11]. Such heterogeneity is known to support cancer cell survival in response to treatment due pre-existent or arising resistant clones. The major challenge remains how best to apply NGS to a single sample in space and time that addresses this well-recognized underlying complexity [12]. New technologies using NGS have lately been applied to liquid biopsy and showed potential to overcome some of the challenges of intra-tumor heterogeneity and further as a viable alternative to acquire material for genomic testing where biopsy or surgery are not an option, in advanced cancers. The sequencing of circulating cell-free DNA (cfDNA) obtained from blood promises a less invasive clinical testing process that has the potential to identify tumor mutations to guide treatment or identify mechanisms of treatment resistance; however, sensitivity remains an issue [13,14]. The American Society of Clinical Oncology and College of American Pathologists indicated that currently there is insufficient evidence for routine clinical use of cfDNA analysis and called for more studies to confirm clinical validity [15].

NGS has also revolutionized the study of the tumor transcriptome by identification of novel transcripts, fusion genes. Whole transcriptome sequencing also allow new insights into the tumor microenviroment using deconvolution methods to estimate the abundance of different cell types from a mixed cell population [16]. NGS has also revealed a tumorigenic role for non-coding RNAs with increasing evidence of their involvement in regulating gene expression affecting cancer cell plasticity and adding an extra layer of complexity in cancer biology. Non-coding RNAs have been described as regulatory molecules that mediate cellular processes, including chromatin remodeling, transcription, post-transcriptional modifications and signal transduction, which have the potential to be therapeutic targets [17]. However, there is still a great amount of research needed if this space to unravel the complex network non-coding RNAs in cancer biology.

In the research setting, new sequencing technologies that rely on the principals of NGS continue to be developed. For example: (1) ATAC-Seq (Assay for Transposase Accessible Chromatin with NGS) an approach for genome-wide chromatin accessibility profiling using hyperactive Tn5 transposase, which simultaneously cuts DNA and inserts sequencing adaptors [18]. (2) ChiP-Seq (chromatin immunoprecipitation combined with NGS) uses antibodies to select specific proteins or nucleosomes, to enrich for DNA fragments that are bound to these proteins or nucleosomes that are recovered and sequenced to provide a map of binding sites [19,20]. These are important tools to identify genome-wide changes in regulatory elements such transcription factor binding sites, enhancer and insulators regulating gene expression.

RNASeq together with single cell sequencing and spatial transcriptomics are bring new information allowing the study of subpopulations of cells and their relationships in the tumor microenvironment leading to unprecedented gain in knowledge in cancer biology [21]. The much anticipated single-molecule long-read sequencing technologies are beginning to be widely available as costs are reduced and platforms delivery stabilizes. Therefore, long-read sequencing is poised to deliver the next layer of new insights in cancer research by allowing the resolution of repeat sequences (a limiting ability of short read NGS), complex structural rearrangements and transcript isoform structure. Overall, the NGS revolution has allowed researchers to ask and answer questions about tumor biology that could not be performed previously. Developments in NGS and emerging novel technologies continue to change the cancer research landscape to advance our knowledge with the hope that in the near future we can further improve our ability to prevent, diagnose and treat cancers.

References

1. Hanahan, D.; Weinberg, R.A. Hallmarks of cancer: The next generation. *Cell* **2011**, *144*, 646–674. [CrossRef] [PubMed]
2. Cancer Genome Atlas Research Network; Weinstein, J.N.; Collisson, E.A.; Mills, G.B.; Shaw, K.R.; Ozenberger, B.A.; Ellrott, K.; Shmulevich, I.; Sander, C.; Stuart, J.M. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat. Genet.* **2013**, *45*, 1113–1120. [CrossRef] [PubMed]
3. Bailey, M.H.; Tokheim, C.; Porta-Pardo, E.; Sengupta, S.; Bertrand, D.; Weerasinghe, A.; Colaprico, A.; Wendl, M.C.; Kim, J.; Reardon, B.; et al. Comprehensive Characterization of Cancer Driver Genes and Mutations. *Cell* **2018**, *173*, 371–385. [CrossRef] [PubMed]
4. International Cancer Genome Consortium; Hudson, T.J.; Anderson, W.; Artez, A.; Barker, A.D.; Bell, C.; Bernabé, R.R.; Bhan, M.K.; Calvo, F.; Eerola, I.; et al. International network of cancer genome projects. *Nature* **2010**, *464*, 993–998.
5. ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium. Pan-cancer analysis of whole genomes. *Nature* **2020**, *578*, 82–93. [CrossRef]
6. Hutter, C.; Zenklusen, J.C. The Cancer Genome Atlas: Creating Lasting Value beyond Its Data. *Cell* **2018**, *173*, 283–285. [CrossRef]
7. Alexandrov, L.B.; Nik-Zainal, S.; Wedge, D.C.; Aparicio, S.A.; Behjati, S.; Biankin, A.V.; Bignell, G.R.; Bolli, N.; Borg, A.; Borresen-Dale, A.L.; et al. Signatures of mutational processes in human cancer. *Nature* **2013**, *500*, 415–421. [CrossRef]
8. Nik-Zainal, S.; Davies, H.; Staaf, J.; Ramakrishna, M.; Glodzik, D.; Zou, X.; Martincorena, I.; Alexandrov, L.B.; Martin, S.; Wedge, D.C.; et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* **2016**, *534*, 47–54. [CrossRef]
9. Dagogo-Jack, I.; Shaw, A.T. Tumor heterogeneity and resistance to cancer therapies. *Nat. Rev. Clin. Oncol.* **2018**, *15*, 81–94. [CrossRef]
10. Nik-Zainal, S.; Van Loo, P.; Wedge, D.C.; Alexandrov, L.B.; Greenman, C.D.; Lau, K.W.; Raine, K.; Jones, D.; Marshall, J.; Ramakrishna, M.; et al. The life history of 21 breast cancers. *Cell* **2012**, *149*, 994–1007. [CrossRef]
11. Gerstung, M.; Jolly, C.; Leshchiner, I.; Dentro, S.C.; Gonzalez, S.; Rosebrock, D.; Mitchell, T.J.; Rubanova, Y.; Anur, P.; Yu, K.; et al. The evolutionary history of 2,658 cancers. *Nature* **2020**, *578*, 122–128. [CrossRef]
12. Yates, L.R.; Campbell, P.J. Evolution of the cancer genome. *Nat. Rev. Genet.* **2012**, *13*, 795–806. [CrossRef] [PubMed]
13. Zill, O.A.; Banks, K.C.; Fairclough, S.R.; Mortimer, S.A.; Vowles, J.V.; Mokhtari, R.; Gandara, D.R.; Mack, P.C.; Odegaard, J.I.; Nagy, R.J.; et al. The Landscape of Actionable Genomic Alterations in Cell-Free Circulating Tumor DNA from 21,807 Advanced Cancer Patients. *Clin. Cancer Res.* **2018**, *24*, 3528–3538. [CrossRef] [PubMed]
14. Li, B.T.; Janku, F.; Jung, B.; Hou, C.; Madwani, K.; Alden, R.; Razavi, P.; Reis-Filho, J.S.; Shen, R.; Isbell, J.M.; et al. Ultra-deep next-generation sequencing of plasma cell-free DNA in patients with advanced lung cancers: Results from the Actionable Genome Consortium. *Ann. Oncol.* **2019**, *30*, 597–603. [CrossRef] [PubMed]
15. Merker, J.D.; Oxnard, G.R.; Compton, C.; Diehn, M.; Hurley, P.; Lazar, A.J.; Lindeman, N.; Lockwood, C.M.; Rai, A.J.; Schilsky, R.L.; et al. Circulating tumor DNA analysis in patients with cancer: American Society of Clinical Oncology and College of American Pathologists Joint Review. *J. Clin. Oncol.* **2018**, *36*, 1631–1641. [CrossRef]
16. Newman, A.M.; Liu, C.L.; Green, M.R.; Gentles, A.J.; Feng, W.; Xu, Y.; Hoang, C.D.; Diehn, M.; Alizadeh, A.A. Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* **2015**, *12*, 453–457. [CrossRef]
17. Anastasiadou, E.; Jacob, L.S.; Slack, F.J. Non-coding RNA networks in cancer. *Nat. Rev. Cancer* **2018**, *18*, 5–18. [CrossRef] [PubMed]
18. Sun, Y.; Miao, N.; Sun, T. Detect accessible chromatin using ATAC-sequencing, from principle to applications. *Hereditas* **2019**, *156*, 29. [CrossRef]
19. Jiang, S.; Mortazavi, A. Integrating ChIP-seq with other functional genomics data. *Brief. Funct. Genomics* **2018**, *17*, 104–115. [CrossRef] [PubMed]

20. Park, P. ChIP–seq: Advantages and challenges of a maturing technology. *Nat. Rev. Genet.* **2009**, *10*, 669–680. [CrossRef]
21. Suvà, M.L.; Tirosh, I. Single-Cell RNA Sequencing in Cancer: Lessons Learned and Emerging Challenges. *Mol. Cell* **2019**, *75*, 7–12. [CrossRef] [PubMed]

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).